



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : G01N 33/53, C12Q 1/00, 1/48, C12P 21/06, C12N 1/20, 5/00		A1	(11) International Publication Number: WO 95/07463
			(43) International Publication Date: 16 March 1995 (16.03.95)
(21) International Application Number: PCT/US94/10165		(74) Agent: WHITE, John, P.; Cooper & Dunham, 30 Rockefeller Drive, East Falmouth, MA 02536 (US).	
(22) International Filing Date: 9 September 1994 (09.09.94)		(81) Designated States: AU, CA, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 08/119,678 10 September 1993 (10.09.93) US 08/192,274 4 February 1994 (04.02.94) US		(81) Designated States: AU, CA, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(60) Parent Applications or Grants (63) Related by Continuation US 08/119,678 (CIP) Filed on 10 September 1993 (10.09.93) US 08/192,274 (CIP) Filed on 4 February 1994 (04.02.94)		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(71) Applicants (for all designated States except US): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; West 116th Street & Broadway, New York, NY 10027 (US). WOODS HOLE OCEANOGRAPHIC INSTITUTE [US/US]; Woods Hole, MA 02543 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): CHALFIE, Martin [US/US]; Apartment 51, 15 Claremont Avenue, New York,			

(54) Title: USES OF GREEN FLUORESCENT PROTEIN

(57) Abstract

This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein. This invention also provides living organisms which comprise the above-described cell. This invention also provides a method for selecting cells expressing a protein of interest which comprises: a) introducing into the cells a DNA molecule having DNA sequence encoding the protein of interest and DNA molecule having DNA sequence encoding a green fluorescent protein; b) culturing the introduced cells under conditions permitting expression of the green fluorescent protein and the protein of interest; and c) selecting the cultured cells which express green fluorescent protein, thereby selecting cells expressing the protein of interest. Finally, this invention provides various uses of a green fluorescent protein.

### USES OF GREEN FLUORESCENT PROTEIN

This application is a continuation-in-part of United States Application Serial Nos. 08/119,678 and 08/192,274, filed September 10, 1993 and February 4, 1994, respectively, the contents of which are hereby incorporated by reference.

The invention disclosed herein was made with Government support under NIH Grant No. 5R01GM30997 from the Department of Health and Human Services. Accordingly, the U.S. Government has certain rights in this invention.

#### Background of the invention

Throughout this application various references are referred to within parenthesis. Disclosures of these publications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains. Full bibliographic citation for these references may be found at the end of this application, preceding the sequence listing and the claims.

Several methods are available to monitor gene activity and protein distribution within cells. These include the formation of fusion proteins with coding sequences for  $\beta$ -galactosidase (22), and luciferases (22). The usefulness of these methods is often limited by the requirement to fix cell preparations or to add exogenous substrates or cofactors. This invention disclose a method of examining gene expression and protein localization in living cells that requires no exogenously-added compounds.

This method uses a cDNA encoding the Green fluorescent

- 2 -

Protein (GFP) from the jelly fish *Aequorea victoria* (3). In *A. victoria*, GFP absorbs energy generated by aequorin upon the stimulation by calcium and emits a green light.

- 5     This invention discloses that GFP expressed in prokaryotic and eukaryotic cells is capable of producing a strong green fluorescence when excited with near UV or blue light. Since this fluorescence requires no additional gene products from *A. victoria*, chromophore  
10    formation is not species specific.

- 3 -

Summary of the invention

5 This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein. This invention also provides living organisms comprising the above-described cell.

10 This invention provides a method for selecting cells expressing a protein of interest which comprises: a) introducing into the cells a DNAI molecule having DNA sequence encoding the protein of interest and DNAII molecule having DNA sequence encoding a green fluorescent  
15 protein; b) culturing the introduced cells in conditions permitting expression of the green fluorescent protein and the protein of interest; and c) selecting the cultured cells which express green fluorescent protein, thereby selecting cells expressing the protein of  
20 interest.

This invention also provides a method for localizing a protein of interest in a cell: a) introducing into a cell  
25 a DNA molecule having DNA sequence encoding the protein of interest linked to DNA sequence encoding the green fluorescent protein such that the protein produced by the DNA molecule will have the protein of interest fused to the green fluorescent protein; b) culturing the cell in conditions permitting expression of the fused protein;  
30 c) detecting the location of the fused protein product, thereby localizing the protein of interest.

Brief Description of Figures

- 5           Figure 1   Expression of GFP in *E. coli*. The bacteria on the right side of the figure have the GFP expression plasmid. This photograph was taken while irradiating the agar plate with a hand-held long-wave UV source.
- 10          Figure 2   Excitation and Emission Spectra of *E. coli*-generated GFP (solid lines) and purified *A. victoria* GFP (L form; dotted lines).
- 15          Figure 3   Expression of GFP in a first stage *Caenorhabditis elegans* larva. Two touch receptor neurons (PLML and ALML) and one other neuron of unknown function (ALNL) are indicated. Processes can be seen projecting from all three cell bodies. The arrow points to the nerve ring branch from the ALML cell (out of focus). The background fluorescence is due to the animal's autofluorescence.
- 20

Detailed Description of the Invention

Throughout this application, the following standard  
5 abbreviations are used to indicate specific nucleotides:

C=cytosine	A=adenosine
T=thymidine	G=guanosine

10 This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein.

15 This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein, wherein the cell is selected from a group  
20 consisting essentially of bacterial cell, yeast cell, fungal cell, insect cell, nematode cell, plant or animal cell.

Suitable animal cells include, but are not limited to  
25 Vero cells, HeLa cells, Cos cells, CV1 cells and various vertebral, invertebral, mammalian cells.

In an embodiment, the bacterial cell is *Escherichia coli*.

30 As used herein, "a regulatory element" from a gene is the DNA sequence which is necessary for the transcription of the gene.

In this invention, the term "operatively linked" means  
35 that following such a link the regulatory element can

- 6 -

direct the transcription of the linked protein-coding DNA sequence.

5 The gene encoding a green fluorescent protein includes DNA molecules coding for polypeptide analogs, fragments or derivatives of antigenic polypeptides which differ from naturally-occurring forms in terms of the identity or location of one or more amino acid residues (deletion analogs containing less than all of the residues  
10 specified for the protein, substitution analogs wherein one or more residues specified are replaced by other residues and addition analogs where in one or more amino acid residues is added to a terminal or medial portion of the polypeptides) and which share some or all properties  
15 of naturally-occurring forms.

These DNA molecules include: the incorporation of codons "preferred" for expression by selected mammalian or non-mammalian hosts; the provision of sites for cleavage by  
20 restriction endonuclease enzymes; and the provision of additional initial, terminal or intermediate DNA sequences that facilitate construction of expression vectors.

25 As an example, plasmid pGFP10.1 codes for a mutated GFP protein having the 80th amino acid residue as an arginine rather than a glutamine predicted to be in native GFP from *A. victoria*. This mutated protein retains the property to fluoresce like the natural protein.

30 In an embodiment, the regulatory element is a promoter. In a further embodiment, the promoter is activated by a heavy metal. Such promoters are well-known in the art (J.H. Freedman, L.W. Slice, A. Fire, and C.S. Rubin  
35 (1993) *Journal of Biological Chemistry*, 268:2554).

- 7 -

In another embodiment, the promoter is that from a cytochrome P450 gene. Cytochrome P450 is well-known in the art and there are a number of P450 promoters known.

5 In still another embodiment, the promoter is that from a stress protein gene. Such stress proteins are well-known in the art (E.G. Stringham, D.K. Dixon, D. Jones and E.D. Candido (1992) *Molecular Biology of the Cell*, 3:221; and William J. Welch (May, 1993), *Scientific American*, page  
10 56). In a further embodiment, the stress protein is a heat-shock protein.

This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a  
15 gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein, wherein the promoter is from a gene necessary for the viability of a cell.

20 In another embodiment, the regulatory element is an enhancer. Enhancers are well-known in the art.

This invention provides a cell comprising a DNA molecule having a regulatory element from a gene, other than a  
25 gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein, wherein the DNA sequence encodes the *Aequorea victoria* green fluorescent protein.

30 In an embodiment, the *Aequorea victoria* green fluorescent protein is cloned in a plasmid. This plasmid is a modification of the pBS(+) (formerly called Bluescribe +) vector (Stratagene®) which has inserted within it an *EcoRI* fragment containing the cDNA sequence of the  
35 *Aequorea victoria* green fluorescent protein (as modified



- 8 -

herein). The fragment was obtained from  $\lambda$ GFP10 (Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J., (1992) Primary structure of the *Aequorea victoria* green fluorescent protein. *Gene*, 111:229-233) by  
5 amplification using the polymerase chain reaction (Saiki, R.K., Gelfand, D.H., Stoffel, S., Sharf, S.J., Higuchi, G.T., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239:487-491) with  
10 primers flanking the *EcoRI* sites and subsequent digestion with *EcoRI*. The sequence of the cDNA in pGFP10.1 differs from the published sequence (5) by a change of the 80th codon of the coding sequence from CAG to CGG, a change that replaces a glutamine with arginine in the protein  
15 sequence.

This pGFP10.1 plasmid was deposited on September 1, 1993 with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A. under  
20 the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganism for the Purposes of Patent Procedure. Plasmid pGFP10.1 was accorded ATCC Accession Number 75547.

25 In another embodiment, this invention provide a bacterial cell which is expressing the green fluorescent protein. In an further embodiment, the bacterial cell is an *E.coli* cell. In a still further embodiment, this *E.coli* cell is designated SMC1 (ATCC Accession No. 69554).

30 This SMC1 bacterial cell was deposited on February 4, 1994, 1993 with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A. under the provisions of the Budapest Treaty for  
35 the International Recognition of the Deposit of

- 9 -

Microorganism for the Purposes of Patent Procedure.  
Bacterial cell SMC1 was accorded ATCC Accession Number  
69554.

5 This invention further provides an isolated green  
fluorescent protein produced from the above-described  
cells which comprise a DNA molecule having a regulatory  
element from a gene, other than a gene encoding a green  
10 fluorescent protein operatively linked to a DNA sequence  
encoding the green fluorescent protein. This isolated  
green fluorescent protein can then be further modified in  
vitro for various uses.

15 This invention disclose an efficient method for  
expression of green fluorescent protein such that large  
amount of the protein could be produced. Methods to  
isolate expressed protein have been well-known and  
therefore, green fluorescent protein may be isolated  
easily.

20 This invention provides a living organism comprising the  
cell comprising a DNA molecule having a regulatory  
element from a gene, other than a gene encoding a green  
fluorescent protein operatively linked to a DNA sequence  
25 encoding the green fluorescent protein.

In another embodiment, the living organism is human. In  
another embodiment, the living organism is a mouse. The  
living organism may be other mammals. In addition, this  
30 invention is applicable to other vertebrates, non-  
vertebrates and living organisms.

In an embodiment, the living organism is *C. elegans*. In  
still another embodiment, the living organism is  
35 *Drosophila*, zebra fish, virus or bacteriophage.

- 10 -

A bacteriophage carrying the green fluorescent protein gene can infect a particular type of bacteria. The infection may be easily detected via the expression of the green fluorescent protein. Therefore, by using appropriate bacteriophages, the presence of that particular type of bacteria may be detected.

Similarly, a eucaryotic virus carrying the green fluorescent protein gene may infect a specific cell type. The infection may be easily detected by monitoring the expression of the green fluorescent protein.

Methods to introduce exogenous genetic material into a cell are well-known in the art. For example, exogenous DNA material may be introduced into the cell by calcium phosphate precipitation technology. Other technologies, such as the retroviral vector technology, electroporation, lipofection and other viral vector systems such as adeno-associated virus system, or microinjection may be used.

The above-described cells and living organisms are useful to detect effects of external stimulus to the regulatory element. The stimulus may have direct or indirect effects on the regulatory element. Such effects will be detectable through either the induction of expression and production of the green fluorescent protein or switching off the expression of the green fluorescent protein.

Cells expressing the green fluorescent proteins may be conveniently separated by a fluorescence-activated cell sorter.

These cells and organisms may be used to detect the presence of different molecules in various kinds of

- 11 -

biological samples such as blood, urine or saliva. By operatively linking a regulatory element of the gene which is affected by the molecule of interest to a green fluorescent protein, the presence of the molecules will affect the regulatory element which in turn will affect the expression of the green fluorescent protein. Therefore, the above-described cells are useful for the detection of molecules. Such detection may be used for diagnostic purposes. An example of such a molecule is a hormone.

This invention provides a living organism comprising the cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein, wherein the regulatory element is for a stress protein.

This invention provides a living organism comprising the cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein, wherein the stress protein is a heat-shock protein.

This invention provides a method to produce green fluorescent protein comprising a) culturing the above-described cells comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein; and b) isolating and purifying the green fluorescent protein produced from the cells. Standard methods for isolating and purifying proteins are well-known in the art. In an embodiment, the cells used for production of

- 12 -

green fluorescent proteins are *E. coli* cells. In a further embodiment, the *E. coli* cells are cultured aerobically.

5 This invention provides a method to synthesize green fluorescent protein comprising a) culturing the cell designated SMC1; and b) isolating and purifying the green fluorescent protein produced from the cell.

10 This invention provides a method for selecting cells expressing a protein of interest which comprises: a) introducing into the cells a DNAI molecule having DNA sequence encoding the protein of interest and DNAII molecule having DNA sequence encoding a green fluorescent  
15 protein; b) culturing the introduced cells in conditions permitting expression of the green fluorescent protein and the protein of interest; and c) selecting the cultured cells which express green fluorescent protein, thereby selecting cells expressing the protein of  
20 interest.

This invention also provides the above method, wherein the cells are selected from a group consisting essentially of bacterial cells, yeast cells, fungal  
25 cells, insect cells, nematode cells, plant or animal cells. Suitable animal cells include, but are not limited to Vero cells, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells.

30 In an embodiment, DNAI and DNAII are linked. In another embodiment, the DNA encodes the *Aequorea victoria* green fluorescent protein.

This invention provides a method for localizing a protein  
35 of interest in a cell which comprises: a) introducing

- 13 -

into a cell a DNA molecule having DNA sequence encoding the protein of interest linked to DNA sequence encoding a green fluorescent protein such that the protein produced by the DNA molecule will have the protein of interest fused to the green fluorescent protein; b) culturing the cell in conditions permitting expression of the fused protein; and c) detecting the fused protein composed of the green fluorescent protein in the cell, thereby localizing a protein of interest in a cell.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and translation initiation sequences for ribosome binding. For example, a bacterial expression vector includes a promoter such as the lac promoter and for translation initiation the Shine-Dalgarno sequence and the start codon ATG. Similarly, a eukaryotic expression vector includes a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon ATG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well-known in the art, for example the methods described above for constructing vectors in general.

To maximize the expression of the green fluorescent protein, the sequence flanking the translation initiation codon may be modified (reviewed by Kozak, 1984), compilation and analysis of sequences upstream from the translation start site in eucaryotic mRNAs. *Nucl. Acids. Res.* 12:857-872). A sequence may then be generated to produce higher amounts of the GFP protein.

In addition, artificial introns may be introduced so as to increase the production of the protein.

- 14 -

Other special targeting sequences may be inserted into the GFP gene. One such targeting sequence is the nuclear localization signal (such as the SV40 nuclear localization signal).

5

The host cell of the above expression system may be selected from the group consisting of the cells where the protein of interest is normally expressed, or foreign cells such as bacterial cells (such as *E. coli*), yeast cells, fungal cells, insect cells (such as Sf9 cell in the baculovirus expression system), nematode cells, plant or animal cells, where the protein of interest is not normally expressed. Suitable animal cells include, but are not limited to Vero cells, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells.

10

15

In an embodiment of the method for localizing a protein of interest in a cell, the DNA encoding the green fluorescent protein is from *Aequorea victoria*.

20

This invention provides a method for localizing a protein of interest in a cell which comprises: a) introducing into a cell a DNA molecule having DNA sequence encoding the protein of interest linked to DNA sequence encoding a green fluorescent protein such that the protein produced by the DNA molecule will have the protein of interest fused to the green fluorescent protein; b) culturing the cell in conditions permitting expression of the fused protein; and c) detecting the location of the fused protein composed of green fluorescent protein in the cell, thereby localizing a protein of interest in a cell, wherein the cell normally expressing the protein of interest.

25

30

This invention provides a method for detecting expression

35

- 15 -

of a gene in a cell which comprises: a) introducing into the cell a DNA molecule having DNA sequence of the gene linked to DNA sequence encoding a green fluorescent protein such that the regulatory element of the gene will control expression of the green fluorescent protein; b) culturing the cell in conditions permitting expression of the gene; and c) detecting the expression of the green fluorescent protein in the cell, thereby indicating the expression of the gene in the cell.

10

This invention provides a method for indicating expression of a gene in a subject which comprises: a) introducing into a cell of the subject a DNA molecule having DNA sequence of the gene linked to DNA sequence encoding a green fluorescent protein such that the regulatory element of the gene will control expression of the green fluorescent protein; b) culturing the cell in conditions permitting expression of the fused protein; and c) detecting the expression of the green fluorescent protein in the cell, thereby indicating the expression of the gene in the cell.

15

20

In an embodiment of the above methods, the green fluorescent protein is the *Aequorea victoria* green fluorescent protein.

25

This invention provides a method for determining the tissue-specificity of transcription of a DNA sequence in a subject which comprises: a) introducing into a cell of the subject a DNA molecule having the DNA sequence linked to DNA sequence encoding a green fluorescent protein such that the DNA sequence will control expression of the green fluorescent protein; b) culturing the subject in conditions permitting the expression of the green fluorescent protein; and c) detecting the expression of

30

35



- 16 -

the green fluorescent protein in different tissue of the subject, thereby determining the tissue-specificity of the expression of the DNA sequence.

5 This invention provides a method for determining the presence of heavy metal in a solution which comprises: a) culturing the cell comprising a DNA molecule having a promoter from a gene, other than a green fluorescent protein operatively linked to a DNA sequence encoding the  
10 green fluorescent protein, wherein transcription at the promoter is activated by a heavy metal in the solution; and b) detecting expression of the green fluorescent protein, the expression of the green fluorescent protein indicates the presence of heavy metal.

15 This invention provides a method for detecting pollutants in a solution which comprises: a) culturing the cell comprising a DNA molecule having a promoter from a gene, other than a green fluorescent protein operatively linked  
20 to a DNA sequence encoding the green fluorescent protein, wherein the promoter is activated by a heavy metal or a toxic organic compound or the promoter is for a stress protein in the solution; and b) detecting expression of the green fluorescent protein, the expression of the  
25 green fluorescent protein indicates the presence of pollutants in the solution.

Finally, this invention provides a method for producing fluorescent molecular weight markers comprising: a)  
30 linking a DNA molecule encoding a green fluorescent protein with a DNA molecule encoding a known amino acid sequence in the same reading frame; b) introducing the linked DNA molecule of step a) in an expression system permitting the expression of a fluorescent protein  
35 encoded by the linked DNA molecule; and c) determining

- 17 -

the molecular weight of the expressed fluorescent protein of step b), thereby producing a fluorescent molecular weight marker.

- 5 Various expression systems are known in the art. The E. coli expression system, one of the commonly used system is described in the following section.

- 10 The determination of molecular weight may be done by comparing the expressed fluorescent protein of step b) with known molecular weight markers. Alternatively, the molecular weight can be predicted by calculation since the linked DNA sequence is known (and so is the amino acid sequence being encoded). In an embodiment, the  
15 expressed fluorescent protein is purified. The purified fluorescent protein can be conveniently used as molecular weight markers.

- 20 This invention will be better understood from the Experimental Details which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

- 18 -

Experimental Details

A cDNA for the *Aequorea victoria* green fluorescent protein (GFP) produces a fluorescent product when  
5 expressed in prokaryotic (*Escherichia coli*) or eukaryotic  
(*Caenorhabditis elegans*) cells. Because exogenous  
substrates and cofactors are not required for this  
fluorescence, GFP expression can be used to monitor gene  
expression and protein localization in living organisms.

10 Light is produced by the bioluminescent jellyfish  
*Aequorea victoria* when calcium binds to the photoprotein  
aequorin (1). Although activation of aequorin in vitro  
or in heterologous cells produces blue light, the  
15 jellyfish produces green light. This latter light is the  
result of a second protein in *A. victoria* that derives  
its excitation energy from aequorin (2), the green  
fluorescent protein (GFP).

20 Purified GFP, a protein of 238 amino acids (3), absorbs  
blue light (maximally at 395 nm with a minor peak at 470  
nm) and emits green light (peak emission at 509 nm with  
a shoulder at 540 nm) (2, 4). This fluorescence is very  
stable; virtually no photobleaching is observed (5).

25 Although the intact protein is needed for fluorescence,  
the same absorption spectral properties found in the  
denatured protein are found in a hexapeptide that starts  
at amino acid 64 (6, 7). The GFP chromophore is derived  
from the primary amino acid sequence through the  
30 cyclization of Ser-dehydroTyr-Gly within this hexapeptide  
(7). The mechanisms that produce the dehydrotyrosine and  
cyclize the polypeptide to form the chromophore are  
unknown. To determine whether additional factors from *A.*  
*victoria* were needed for the production of the  
35 fluorescent protein, applicants tested GFP fluorescence

- 19 -

in heterologous systems. Here applicants show that GFP expressed in prokaryotic and eukaryotic cells is capable of producing a strong green fluorescence when excited by blue light. Because this fluorescence requires no additional gene products from *A. victoria*, chromophore formation is not species specific and occurs either through the use of ubiquitous cellular components or by autocatalysis.

Expression of GFP in *Escherichia coli* (8) under the control of the T7 promoter results in a readily detected green fluorescence (9) that is not observed in control bacteria. Upon illumination with a long-wave UV source, fluorescent bacteria were detected on agar plates containing the inducer isopropyl- $\beta$ -D-thiogalactoside (IPTG) (Fig. 1). When GFP was partially purified from this strain (10), it was found to have fluorescence excitation and emission spectra indistinguishable from those of the purified native protein (Fig. 2). The spectral properties of the recombinant GFP suggest that the chromophore can form in the absence of other *A. victoria* products.

Transformation of the nematode *Caenorhabditis elegans* also resulted in the production of fluorescent GFP (11) (Fig. 3). GFP was expressed in a small number of neurons under the control of a promoter for the *mec-7* gene. The *mec-7* gene encodes a  $\beta$ -tubulin (12) that is abundant in six touch receptor neurons in *C. elegans* and less abundant in a few other neurons (13, 14). The pattern of expression of GFP was similar to that detected by MEC-7 antibody or from *mec-7lacZ* fusions (13-15). The strongest fluorescence was seen in the cell bodies of the four embryonically-derived touch receptor neurons (ALML, ALMR, PLML, PLMR) in younger larvae. The processes from

- 20 -

these cells, including their terminal branches, were often visible in larval animals. In some newly hatched animals, the PLM processes were short and ended in what appeared to be prominent growth cones. In older larvae, the cell bodies of the remaining touch cells (AVM and PVM) were also seen; the processes of these cells were more difficult to detect. These postembryonically-derived cells arise during the first of the four larval stages (16), but their outgrowth occurs in the following larval stages (17), with the cells becoming functional during the fourth larval stage (18). GFP's fluorescence in these cells is consistent with these previous results: no fluorescence was detected in these cells in newly hatched or late first-stage larvae, but it was seen in four of ten late second-stage larvae, all nine early fourth-stage larvae, and seven of eight young adults (19). In addition, moderate to weak fluorescence was seen in a few other neurons (Fig. 3) (20). The details of the expression pattern are being examined.

Like the native protein, GFP expressed in both *E. coli* and *C. elegans* is quite stable (lasting at least ten minutes) when illuminated with 450-490 nm light. Some photobleaching occurs, however, when the cells are illuminated with 340-390 nm or 395-440 nm light (21).

Several methods are available to monitor gene activity and protein distribution within cells. These include the formation of fusion proteins with coding sequences for  $\beta$ -galactosidase, firefly luciferase, and bacterial luciferase (22). Because such methods require exogenously-added substrates or cofactors, they are of limited use with living tissue. Because the detection of intracellular GFP requires only irradiation by near UV or blue light, it is not substrate limited. Thus, it should

- 21 -

provide an excellent means for monitoring gene expression and protein localization in living cells (23, 24). Because it does not appear to interfere with cell growth and function, GFP should also be a convenient indicator of transformation and one that could allow cells to be separated using fluorescence-activated cell sorting. Applicants also envision that GFP can be used as a vital marker so that cell growth (for example, the elaboration of neuronal processes) and movement can be followed in situ, especially in animals that are essentially transparent like *C. elegans* and zebrafish. The relatively small size of the protein may facilitate its diffusion throughout the cytoplasm of extensively branched cells like neurons and glia. Since the GFP fluorescence persists after treatment with formaldehyde (9), fixed preparations can also be examined. In addition, absorption of appropriate laser light by GFP-expressing cells (as has been done for lucifer yellow-containing cells) (25), could result in the selective killing of the cells.

#### Further Experiments on GFP Expression

The TU#58 plasmid, which contains the green fluorescent protein (GFP) coding sequence in the pET3a expression vector (29) was transformed into *Escherichia coli* strain BLR (DE3) (A. Roca, University of Wisconsin: cited in the Novogen Catalogue) using procedures described previously (29). The resulting strain (SMC3), because of the reduced recombination of the host, was much more stable for GFP expression (all the colonies on plates with ampicillin but without the IPTG inducer (29) were brightly fluorescent when viewed with a hand-held UV lamp).

35

- 22 -

A second construct (TU#147), similar to TU#58, was made with pET11 (A.H. Rosenberg, et al. 1987). Expression in BLR (DE3) from this plasmid was more tightly controlled; expression was seen soon after IPTG was added, but only  
5 after some time without inducer.

The SMC3 strain was used to test the requirement for aerobic growth of the bacteria for the production of a fluorescent product. Plates were grown under anaerobic  
10 conditions in a Gas-Pak container according to the instructions of the manufacturer (Becton Dickinson Microbiology Systems). Colony growth was slowed under anaerobic conditions and the resulting colonies were not detectably fluorescent after at least 3 days of growth  
15 under anaerobic conditions (using the hand-held UV lab). Colonies, however, became fluorescent after a day's exposure to room air (some fluorescence was seen after a few hours).

20

## REFERENCES AND NOTES

1. O. Shimomura, F.H. Johnson, Y. Saiga, *J. Cell. Comp. Physiol.* 59, 223 (1962).  
5
2. J. G. Morin and J. W. Hastings, *J. Cell. Physiol.* 77, 313 (1971); H. Morise, O. Shimomura, F. H. Johnson, J. Winant, *Biochemistry* 13, 2656 (1974).
- 10 3. D. C. Prasher, V. K. Eckenrode, W.W. Ward, F. G. Prendergast, M. J. Cormier, *Gene* 111, 229 (1992).
4. W. W. Ward, C. W. Cody, R. C. Hart, M. J. Cormier, *Photochem. Photobiol.* 31, 611 (1980).  
15
5. F. G. Prendergast, personal communication.
6. O. Shimomura, *FEBS Lett.* 104, 220 (1979).
- 20 7. C. W. Cody, D. C. Prasher, W. M. Westler, F. G. Prendergast, W. W. Ward, *Biochemistry* 32, 1212 (1993).
- 25 8. Plasmid pGFP10.1 contains the *EcoRI* fragment encoding the GFP cDNA from  $\lambda$ gfp10 (3) in pBS(+) (Stratagene®). The fragment was obtained by amplification with the polymerase chain reaction [PCR; R. K. Saiki et al., *Science* 239, 487 (1988)] with primers flanking the *EcoRI* sites and subsequent digestion with *EcoRI*. DNA was prepared by the Magic Minipreps procedure (Promega) and sequenced (after an additional ethanol precipitation) on an Applied Biosystems DNA Sequencer 370A at the DNA Sequencing facility at Columbia College of Physicians and Surgeons. The sequence of the cDNA in pGFP10.1  
30  
35



- 24 -

differs from the published sequence by a change in codon 80 within the coding sequence from CAG to CGG, a change that replaces a glutamine residue with arginine [R. Heim, S. Emr, and R. Tsien (personal communication) first alerted us to a possible sequence change in this clone and independently noted the same change.] This replacement has no detectable effect on the spectral properties of the protein (Fig. 2).

An *E. coli* expression construct was made with PCR to generate a fragment with an *NheI* site at the start of translation and an *EcoRI* site 5' to the termination signal of the GFP coding sequence from pGFP10.1. The 5' primer was ACAAGGCTAGCAAAGGAGAAGAAC (Sequence ID No. 1) and the 3' primer was the T3 primer (Stratagene®). The *NheI*-*EcoRI* fragment was ligated into the similarly cut vector pET3a [A.H. Rosenberg et al., Gene 56, 125 (1987)] by standard methods (26). The resulting coding sequence substitutes an Ala for the initial GFP Met, which becomes the second amino acid in the polypeptide. The *E. coli* strain BL21(DE3)Lys S [F. W. Studier and B. A. Moffat, *J. Mol. Biol.* 189, 113 (1986)] was transformed with the resulting plasmid (TU#58) and grown at 37°C. Control bacteria were transformed with pET3a. Bacteria were grown on nutrient plates containing ampicillin (100 µg/ml) and 0.8 mM IPTG. Transformed bacteria from this transformation show green fluorescence when irradiated with ultraviolet light. A recombinant plasmid of this bacteria was used for the experiments described here and the experiment in Figure 2 and the experiment in Note 10. Several months later, applicants noticed that the bacterial

- 25 -

colonies can be divided into two groups: 1) strongly fluorescent; and 2) weakly fluorescent (applicants believe that the weakly fluorescent may have mutated, disabled or partial or completely deleted TU#58). One strongly fluorescent colony was picked to generate the bacterial strain SMC1 (ATCC Accession No. 69554). [A similar PCR-generated fragment (see note 11) was used in applicants' *C. elegans* construct. As others are beginning to use pGFP10.1, applicants have heard that while similar PCR fragments produce a fluorescent product in other organisms (R. Heim, S. Emr, and R. Tsien, personal communication; S. Wang and T. Hazelrigg, personal communication; L. Lanini and F. McKeon, personal communication; see note 23), the *EcoRI* fragment does not (R. Heim, S. Emr, and R. Tsien, personal communication; A. Coxon, J. R. Chaillet, and T. Bestor, personal communication). These results may indicate that elements at the 5' end of the sequence or at the start of translation inhibit expression.]

9. Applicants used a variety of microscopes (Zeiss Axiophot, Nikon Microphot FXA, and Olympus BH2-RFC and BX50) equipped for epifluorescence microscopy. Usually filter sets for fluorescein isothiocyanate fluorescence were used (for example, the Zeiss filter set used a BP450-490 excitation filter, 510 nm dichroic, and either a BP515-565 or a LP520 emission filter), although for some experiments filter sets that excited at lower wavelengths were used (for example, a Zeiss filter set with BP395-440 and LP470 filters and a 460 nm dichroic or with BP340-390 and LP400 filters with a 395 nm dichroic). In some instances a xenon lamp appeared to give a more intense fluorescence than a mercury lamp when

- 26 -

5 cells were illuminated with light around 470 nm, although usually the results were comparable. No other attempts were made to enhance the signal (for example, by using low intensity light cameras), although this may be useful in some instances.

10 Previous experiments had shown that the native protein was fluorescent after glutaraldehyde fixation (W. W. Ward, unpub. data). S. Wang and T. Hazelrigg (personal communication; 23) have found that GFP fusion proteins in *Drosophila melanogaster* are fluorescent after formaldehyde fixation. Applicants have confirmed that fluorescence persists after formaldehyde fixation with applicants' *C.*  
15 *elegans* animals and with recombinant GFP isolated from *E. coli*. The chemicals in nail polish, which is often used to seal cover slips, however, did appear to interfere with the *C. elegans* GFP fluorescence.

20 10. In the applicants' initial experiments, GFP was purified from 250 ml cultures of BL21(DE3)Lys S bacteria containing TU#58; bacteria were grown in LB broth (26) containing ampicillin (100 µg/ml) and 0.8  
25 mM IPTG. Induction was best when IPTG was present continually. Nevertheless, subsequent experiments with bacterial strain SMC1 indicate that the bacteria could not grow in the constant presence of IPTG but can be induced by the IPTG during the log  
30 phase growth. The production of fluorescent protein is best at room temperature. Cells were washed in 4 ml of 10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM MgCl<sub>2</sub>, and 10 mM dithiothreitol [A. Kumagai and W. G. Dunphy, *Cell* 64, 903 (1991)] and then sonicated  
35 (2 X 20 sec) in 4 ml of the same buffer containing

- 27 -

0.1 mM PMSF, pepstatin A (1  $\mu$ g/ml), leupeptin (1  $\mu$ g/ml), and aprotinin (2  $\mu$ g/ml), and centrifuged at 5,000 rpm for 5 min in the cold. The supernatant was centrifuged a second time (15,000 rpm for 15 min) and then diluted sevenfold with 10 mM Tris (pH 8.0), 10 mM EDTA, and 0.02% NaN<sub>3</sub>. Corrected excitation and emission spectra were obtained with a SPEX F1T11 spectrofluorometer and compared with the purified L isoprotein form of GFP from *A. victoria* (M. Cutler, A. Roth, and W. W. Ward, unpub. data). The excitation spectra were measured from 300 - 500 nm with a fixed emission wavelength of 509 nm, and the emission spectra were measured from 410 - 600 nm with a fixed excitation of 395 nm. All spectra were recorded as signal-reference data (where the reference is a direct measurement of the lamp intensity with a separate photomultiplier tube) at room temperature with 1 sec integration times and 1 nm increments. The spectral band widths were adjusted to 0.94 nm for all spectra.

11. Wild-type and mutant *C. elegans* animals were grown and genetic strains were constructed according to S. Brenner, *Genetics* 77, 71 (1974).

The plasmid pGFP10.1 was used as a template for PCR (with the 5' primer GAATAAAAGCTAGCAAAGATGAGTAAAG (Sequence ID No. 2) and the 3' T3 primer) to generate a fragment with a 5' NheI site (at the start of translation) and a 3' EcoRI site (3' of the termination codon). The DNA was cut to produce an NheI - EcoRI fragment that was ligated into plasmid pPD 16.51 (12, 27), a vector containing the promoter of the *C. elegans* *mec-7* gene. Wild-type *C. elegans* were transformed by coinjecting this DNA (TU#64) and

- 28 -

the DNA for plasmid pRF4, which contains the dominant *rol-6(su1006)* mutation, into adult *C. elegans* gonads as described by C. M. Mello, J. M. Kramer, D. Stinchcomb, and V. Ambros, *EMBO J.* 10, 3959 (1991). A relatively stable line was isolated (TU1710) and the DNA it carried was integrated as described by Mitani et al. (15) to produce the integrated elements *uIs3* and *uIs4* (in strains TU1754 and TU1755, respectively).

Living *C. elegans* animals were mounted on agar (or agarose) pads as described (16), often with 10 mM NaN<sub>3</sub> as an anesthetic (28) (another nematode anesthetic, phenoxypropanol, quenched the fluorescence) and examined with either a Zeiss universal or axiophot microscope. For *C. elegans*, a long-pass emission filter works best because the animal's intestinal autofluorescence, (which increases as the animal matures), appears yellow (with band-pass filters the autofluorescence appears green and obscures the GFP fluorescence).

Because much more intense fluorescence was seen in *uIs4* than *uIs3* animals (for example, it was often difficult to see the processes of the ALM and PLM cells in *uIs3* animals when the animals were illuminated with a mercury lamp), the former have been used for the observations reported here. The general pattern of cell body fluorescence was the same in both strains and in the parental, nonintegrated strain (fluorescence in this strain was as strong as that in the *uIs4* animals). The *uIs4* animals, however, did show an unusual phenotype: both the ALM and PLM touch cells were often displaced anteriorly. The mature cells

- 29 -

- usually had processes in the correct positions, although occasional cells had abnormally-projecting processes. These cells could be identified as touch receptor cells, because the fluorescence was dependent on *mec-3*, a homeobox gene that specifies touch cell fate (13, 15, 18, 28). *mec-7* expression is reduced in the ALM touch cells of the head (but not as dramatically in the PLM touch cells of the tail) in *mec-3* gene mutants (13, 15). Applicants find a similar change of GFP expression in a *mec-3* mutant background for both *uIs3* and *uIs4*. Thus, GFP accurately represents the expression pattern of the *mec-7* gene. It is likely that the reduced staining in *uIs3* animals and the misplaced cells in *uIs4* animals is the result of either secondary mutations or the amount and position of the integrated DNA.
12. C. Savage, M. Hamelin, J. G. Culotti, A. Coulson, D. G. Albertson, M. Chalfie, *Genes Dev.* 3, 870 (1989).
  13. M. Hamelin, I. M. Scott, J. C. Way, J. G. Culotti, *EMBO J.* 11, 2885 (1992).
  14. A. Duggan and M. Chalfie, unpub. data.
  15. S. Mitani, H. P. Du, D. H. Hall, M. Driscoll, M. Chalfie, *Development* 119, 773 (1993).
  16. J. E. Sulston and H. R. Horvitz, *Develop. Biol.* 56, 110 (1977).
  17. W. W. Walthall and M. Chalfie, *Science* 239, 643 (1988).
  18. M. Chalfie and J. Sulston, *Dev. Biol.* 82, 358

- 30 -

(1981).

19. In adults, the thicker size of the animals and the more intense autofluorescence of the intestine tend to obscure these cells.
20. These include several cells in the head (including the FLP cells) and tail of newly hatched animals and the BDU cells, a pair of neurons just posterior to the pharynx. Expression of *mec-7* in these cells has been seen previously (13, 15). The strongest staining of these non-touch receptor neurons are a pair of cells in the tail that have anteriorly directed processes that project along the dorsal muscle line. It is likely that these are the ALN cells, the sister cells to the PLM touch cells [J. G. White, E. Southgate, J. N. Thomson, and S. Brenner, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 314, 1 (1986).]
21. The photobleaching with 395-440 nm light is further accelerated, to within seconds, in the presence of 10 mM  $\text{NaN}_3$ , which is used as a *C. elegans* anesthetic (11). However, when cells in *C. elegans* have been photobleached, some recovery is seen within 10 min. Further investigation is needed to determine whether this recovery represents de novo synthesis of GFP. Rapid photobleaching (complete within a minute) of the green product was also seen when *C. elegans* was illuminated with 340-390 nm light. Unlike the photobleaching with 395-440 nm light, which abolished fluorescence produced by the 340-390 or 450-490 nm light, photobleaching with 340-390 nm light did not appear to affect the fluorescence produced by 395-490 or 450-490 nm light. Indeed,

- 31 -

- the fluorescence produced by 450-490 nm light appeared to be more intense after brief photobleaching by 340-390 nm light. This selective photobleaching may indicate the production of more than one fluorescent product in the animal. These data on GFP fluorescence within *E. coli* and *C. elegans* is in contrast to preliminary studies that suggest that the isolated native and *E. coli* proteins are very photostable. Applicants do not know whether this in vivo sensitivity to photobleaching is a normal feature of the jellyfish protein (the fluorescence in *A. victoria* has not been examined) or results from the absence of a necessary posttranslational modification unique to *A. victoria* or nonspecific damage within the cells.
22. Reviewed in T. J. Silhavy and J. R. Beckwith, *Microbiol. Rev.* 49, 398 (1985); S. J. Gould and S. Subramani, *Anal. Biochem.* 175, 5 (1988); and G. S. A. B. Stewart and P. Williams, *J. Gen. Microbiol.* 138, 1289 (1992).
23. R. Heim, S. Emr, and R. Tsien (personal communication) have found that GFP expression in *Saccharomyces cerevisiae* can make the cells strongly fluorescent without causing toxicity. S. Wang and T. Hazelrigg (personal communication) have found that both C-terminal and N-terminal protein fusions with GFP are fluorescent in *Drosophila melanogaster*. L. Lanini and F. McKeon (personal communication) have expressed a GFP protein fusion in mammalian (COS) cells. E. Macagno (personal communication) is expressing GFP in leeches. T. Hughes (personal communication) is expressing GFP in mammalian HEK293 cells.



- 32 -

24. Applicants have generated several other plasmid constructions that may be useful to investigators. These include a pBluescript II KS (+) derivative (TU#65) containing a KpnI - EcoRI fragment encoding GFP with an AgeI site 5' to the translation start and a BsmI site at the termination codon. Also available are gfp versions (TU#60 - TU#63) of the four *C. elegans* lacZ expression vectors (pPD16.43, pPD21.28, pPD22.04, and pPD22.11, respectively) described by Fire et al., 1990 (27) except that they lack the KpnI fragment containing the SV40 nuclear localization signal.
25. J. P. Miller and A. Selverston, *Science* 206, 702 (1979).
26. J. Sambrook, E. F. Fritsch, and T. Maniatis, *Molecular cloning: A laboratory manual*, 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, (1989).
27. A. Fire, S. W. Harrison, and D. Dixon, *Gene* 93, 189 (1990).
28. J. C. Way and M. Chalfie, *Cell* 54, 5 (1988).
29. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., *Science* 263, 802 (1994).

- 33 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

(i) APPLICANT: The Trustees of Columbia University in the City of  
New York and Woods Hole Oceanographic Institute

10

(ii) TITLE OF INVENTION: USES OF GREEN FLUORESCENT PROTEIN

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

15

(A) ADDRESSEE: Cooper & Dunham  
(B) STREET: 30 Rockefeller Plaza  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: United States of America  
(F) ZIP: 10112

20

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

30

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: White, John P.  
(B) REGISTRATION NUMBER: 28,678  
(C) REFERENCE/DOCKET NUMBER: 0575/43557-B-PCT

35

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212) 977-9550  
(B) TELEFAX: (212) 664-0525  
(C) TELEX: 422523 COOP UI

40

## (2) INFORMATION FOR SEQ ID NO:1:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

55

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACAAAGGCTA GCAAAGGAGA AGAAC

25

- 34 -

## (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:  
5 (A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
10 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:  
15 (A) ORGANISM: Escherichia coli

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
20 GAATAAAAGC TAGCAAAGAT GAGTAAAG

28

- 35 -

What is claimed is:

1. A cell comprising a DNA molecule having a regulatory element from a gene, other than a gene encoding a green fluorescent protein operatively linked to a DNA sequence encoding the green fluorescent protein.
2. A cell of claim 1, wherein the cell is selected from a group consisting essentially of bacterial cell, yeast cell, fungal cell, insect cell, nematode cell, plant or animal cell.
3. A cell of claim 1, wherein the regulatory element is a promoter.
4. A cell of claim 3, wherein the promoter is activated by a heavy metal.
5. A cell of claim 3, wherein the promoter is that from a P450 gene.
6. A cell of claim 3, wherein the promoter is from a gene encoding a stress protein.
7. A cell of claim 6, wherein the stress protein is a heat-shock protein.
8. A cell of claim 3, wherein the promoter is from a gene required for cell viability.
9. A cell of claim 1, wherein the regulatory element is an enhancer.
10. A cell of claim 1, wherein the DNA sequence encodes the *Aequorea victoria* green fluorescent protein.

- 36 -

11. A cell of claim 1, wherein the cell is an E. coli cell.
12. The cell of claim 11 designated SMC1 (ATCC Accession No. 69554).
13. An isolated green-fluorescent protein from the cell of claim 1.
14. A living organism comprising the cell of claim 1.
15. A living organism of claim 10, wherein the living organism is *Caenorhabditis elegans*.
16. A living organism of claim 15, wherein the regulatory element is for a stress protein.
17. A living organism of claim 16, wherein the stress protein is a heat-shock protein.
18. A method to produce green fluorescent protein comprising:
  - a) culturing the cell of claim 1; and
  - b) isolating and purifying the green fluorescent protein produced from the cell.
19. A method of claim 18, wherein the cell is an E. coli cell.
20. A method of claim 19, wherein the cell is cultured aerobically.
21. A method of claim 18 wherein the cell is designated

- 37 -

SMC1 (ATCC Accession No. 69554).

22. A method for selecting cells expressing a protein of interest which comprises:

5

a) introducing into the cells a DNAI molecule having DNA sequence encoding the protein of interest and DNAII molecule having DNA sequence encoding a green fluorescent protein;

10

b) culturing the introduced cells under conditions permitting expression of the green fluorescent protein and the protein of interest; and

15

c) selecting the cultured cells which express green fluorescent protein, thereby selecting cells expressing the protein of interest.

20

23. A method of claim 22, wherein DNAI and DNAII are linked.

25

24. A method of claim 22, wherein the cells are selected from a group consisting essentially of bacterial cells, yeast cells, fungal cells, insect cells, nematode cells, plant or animal cells.

30

25. A method of claim 22, wherein the DNAII encodes the *Aequorea victoria* green fluorescent protein.

35

26. A method for localizing a protein of interest in a cell which comprises:

a) introducing into a cell a DNA molecule having DNA sequence encoding the protein of interest linked to DNA sequence encoding a green

- 38 -

fluorescent protein such that the protein produced by the DNA molecule will have the protein of interest fused to the green fluorescent protein;

5

b) culturing the cell under conditions permitting expression of the fused protein; and

10

c) detecting the location of the fused protein which composes the green fluorescent protein in the cell, thereby localizing a protein of interest in a cell.

15

27. A method of claim 26, wherein the cell normally expresses the protein of interest.

28. A method of claim 26, wherein the DNA encodes the *Aequorea victoria* green fluorescent protein.

20

29. A method for detecting expression of a gene in a cell which comprises:

25

a) introducing into the cell a DNA molecule having DNA sequence of the gene linked to DNA sequence encoding a green fluorescent protein such that the regulatory element of the gene will control expression of the green fluorescent protein;

30

b) culturing the cell in conditions permitting expression of the gene; and

35

c) detecting the expression of the green fluorescent protein in the cell, thereby indicating the expression of the gene in the cell.

- 39 -

30. A method for detecting expression of a gene in a subject which comprises:

5 a) introducing into a cell of the subject a DNA molecule having DNA sequence of the gene linked to DNA sequence encoding a green fluorescent protein such that the regulatory element of the gene will control expression of the green fluorescent protein;

10 b) culturing the cell under conditions permitting expression of the fused protein; and

15 c) detecting the expression of the green fluorescent protein in the cell, thereby indicating the expression of the gene in the cell.

20 31. A method of claims 29 or 30, wherein the DNA encodes the *Aequorea victoria* green fluorescent protein.

25 32. A method for determining the tissue-specificity of the transcription of a DNA sequence in a subject which comprises:

30 a) introducing into a cell of the subject a DNA molecule having the DNA sequence linked to a DNA sequence encoding a green fluorescent protein such that the DNA sequence gene will control expression of the green fluorescent protein in the subject;

35 b) culturing the subject in conditions permitting the expression of the green fluorescent protein; and



- 40 -

- c) detecting the expression of the green fluorescent protein in different tissues of the subject, thereby determining the tissue-specificity of the DNA sequence.

5

33. A method for detecting heavy metal in a solution which comprises:

10

- a) culturing the cell of claims 4 or 5 in the solution; and

15

- b) detecting expression of the green fluorescent protein, the expression of the green fluorescent protein indicates the presence of heavy metal.

34. A method for detecting pollutants in a solution which comprises:

20

- a) culturing the cell of claims 4, 5, 6 or 7 in the solution; and

25

- b) detecting expression of the green fluorescent protein, the expression of the green fluorescent protein indicates the presence of a pollutant.

35. A method for producing fluorescent molecular weight markers comprising:

30

- a) linking a DNA molecule encoding a green fluorescent protein with a DNA molecule encoding a known amino acid sequence in the same reading frame;

35

- 41 -

b) introducing the linked DNA molecule of step a) in an expression system permitting the expression of a fluorescent protein encoded by the linked DNA molecule; and

5

c) determining the molecular weight of the expressed fluorescent protein of step b), thereby producing a fluorescent protein molecular weight marker.

10

36. A method of claim 35, further comprising purification of the expressed protein.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US94/10165

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : G01N 33/53; C12Q 1/00, 1/48; C12P 21/06; C12N 1/20, 5/00  
US CL : 435/7.6, 7.9, 15, 69.1, 69.8, 240.2, 252.3, 252.33

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.6, 7.9, 15, 69.1, 69.8, 240.2, 252.3, 252.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAS ONLINE, STN, BIOSIS

search terms: green fluorescent protein?, aequorea, pgfp10.1, nematode?, metallothionein

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GENE, Volume 111, issued 1992, Prasher et al., "Primary structure of the <i>Aequorea victoria</i> green-fluorescent protein", pages 229-233, see especially Fig. 2.	1-36
Y	D.M. Glover, "Gene Cloning. The Mechanics of DNA Manipulation" published 1984 by Chapman and Hall (New York), pages 1-20 and 179-219, see entire document.	1-36

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

18 DECEMBER 1994

Date of mailing of the international search report

04 JAN 1995

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

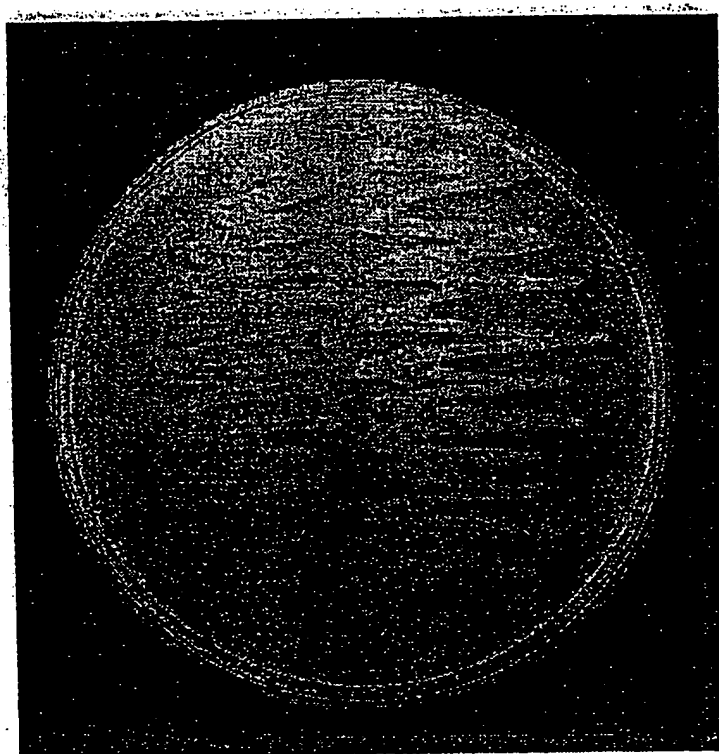
Authorized officer

Hyosuk Kim

Telephone No. (703) 308-0196

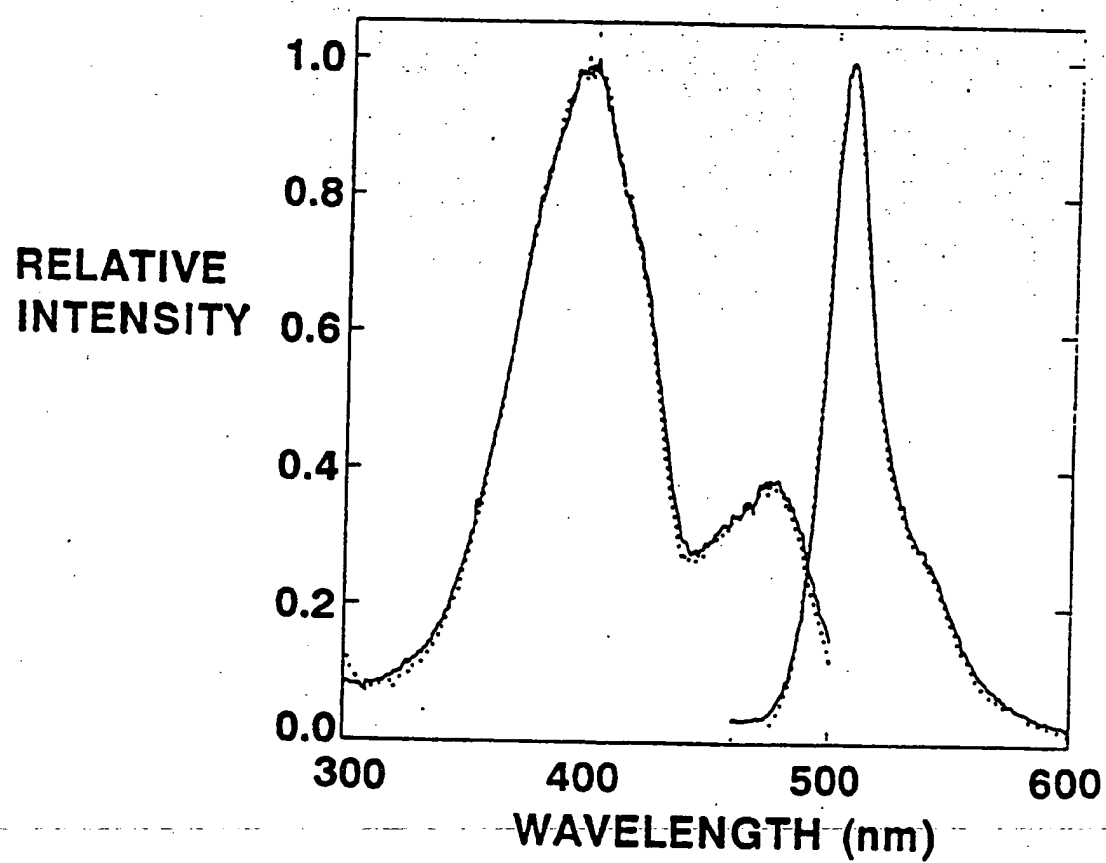
1/3

FIGURE 1



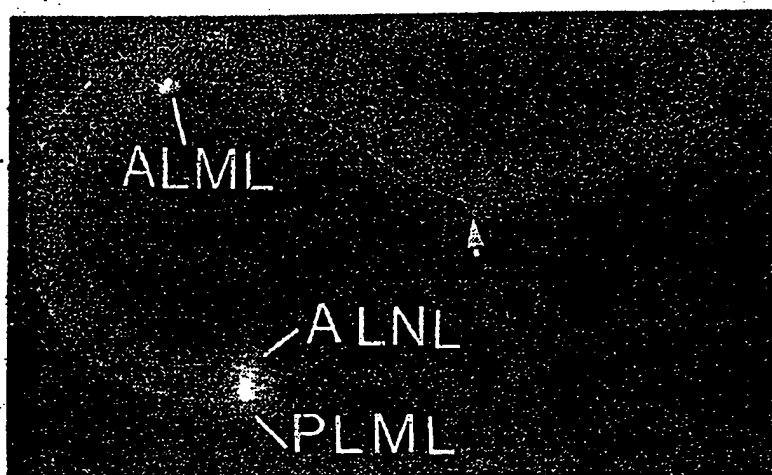
2/3

FIGURE 2



3/3

FIGURE 3



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				